

Rotor Blade Pitch Imbalance Fault Detection for Variable-Speed Marine Current Turbines via Generator Power Signal Analysis

Yufei Tang ¹

¹Florida Atlantic University

October 30, 2023

Abstract

Marine hydrokinetic (MHK) turbines extract renewable energy from oceanic environments. However, due to the harsh conditions that these turbines operate in, system performance naturally degrades over time. Thus, ensuring efficient condition-based maintenance is imperative towards guaranteeing reliable operation and reduced costs for hydroelectric power.

This paper proposes a novel framework aimed at identifying and classifying the severity of rotor blade pitch imbalance faults experienced by marine current turbines (MCTs). In the framework, a Continuous Morlet Wavelet Transform (CMWT) is first utilized to acquire the wavelet coefficients encompassed within the 1P frequency range of the turbine's rotor shaft. From these coefficients, several statistical indices are tabulated into a six-dimensional feature space. Next, Principle Component Analysis (PCA) is employed on the resulting feature space for dimensionality reduction, followed by the application of a K-Nearest Neighbor (KNN) machine learning algorithm for fault detection and severity classification. The framework's effectiveness is validated using a high-fidelity MCT numerical simulation platform, where results demonstrate that pitch imbalance faults can be accurately detected 100% of the time and classified based upon severity more than 97% of the time.

Rotor Blade Pitch Imbalance Fault Detection for Variable-Speed Marine Current Turbines via Generator Power Signal Analysis

Brittney Freeman, Yufei Tang, *Member, IEEE*, Yu Huang, and James VanZwieten

Abstract—Marine hydrokinetic (MHK) turbines extract renewable energy from oceanic environments. However, due to the harsh conditions that these turbines operate in, system performance naturally degrades over time. Thus, ensuring efficient condition-based maintenance is imperative towards guaranteeing reliable operation and reduced costs for hydroelectric power. This paper proposes a novel framework aimed at identifying and classifying the severity of rotor blade pitch imbalance faults experienced by marine current turbines (MCTs). In the framework, a Continuous Morlet Wavelet Transform (CMWT) is first utilized to acquire the wavelet coefficients encompassed within the 1P frequency range of the turbine's rotor shaft. From these coefficients, several statistical indices are tabulated into a six-dimensional feature space. Next, Principle Component Analysis (PCA) is employed on the resulting feature space for dimensionality reduction, followed by the application of a K-Nearest Neighbor (KNN) machine learning algorithm for fault detection and severity classification. The framework's effectiveness is validated using a high-fidelity MCT numerical simulation platform, where results demonstrate that pitch imbalance faults can be accurately detected 100% of the time and classified based upon severity more than 97% of the time.

Index Terms—Marine current turbine, rotor blade imbalance fault, non-intrusive techniques, fault detection and identification, wavelet transform, feature representation

I. INTRODUCTION

THE kinetic energy housed within open-ocean, tidal, and riverine currents represents a highly concentrated source of energy that is both renewable and sustainable. It has been estimated that the energy extracted from the U.S. Gulf Stream current alone has the potential to generate up to 18.6 GW (163 TWh/yr) of electrical power [1]. In addition, the temporally averaged power density encompassed within the section of the Gulf Stream that runs between South Florida and North Carolina reaches 3.3 kW/m^2 [2], while the technically extractable riverine and tidal power encompassed within the entirety of the U.S. has been calculated at 14 GW (123 TWh/yr) [3] and 50 GW (438 TWh/yr) [4] respectively. Various forms of marine current turbines (MCTs) are utilized to convert the kinetic

energy residing within open-ocean, tidal, and riverine currents into electricity. However, for the purposes of this research, only horizontal axis MCTs equipped with permanent magnetic synchronous generators are considered.

There are significant cost reductions that need to be made before marine energy technologies can become cost competitive with other forms of contemporary energy generation (e.g., coal, gas, oil, and nuclear power) [5]. A fundamental challenge impeding the wider scaled implementation of marine-based electricity generation revolves around its currently high Levelized Cost of Energy (LCOE). From the perspective of a marine energy device, the LCOE is an integrated metric that takes into account both the cost and estimated electricity producing capability of the device. Such metrics aid in estimating the revenue per megawatt-hour (MWH) of electricity generation needed to satisfy a minimum rate of return, while also considering the costs associated with building, deploying and operating the device over its lifetime. Investments in marine-based electricity generation are expected to remain sparse until more reliable baseline cost scenarios that utilize standardized cost reporting methodologies and assumptions are made available [6].

Research aimed at reducing the LCOE generally progress down one of two paths. Path one focuses on increasing the marine energy device's overall power output, while path two deals with decreasing the device's operation and maintenance costs (O&M). Regardless of path however, the overall goal is to maximize the amount of revenue obtainable from the grid tied electricity produced by the device.

Path one research was performed in [7], where control strategies for MCTs operating at overrated current speeds were investigated. The results of the research revealed that a flux-weakening control strategy could be used to accelerate the turbine's permanent magnetic synchronous generator (PMSG) over its nominal speed for better power limiting control at high marine current speeds. In one of our recent works [8], an adaptive super-twisting sliding mode control strategy was developed and validated on the same 720-kW numerical simulation platform utilized in this work. Path two research was performed in [9], where a Smart Vibration Monitoring System (SVMS) was developed to improve ocean current turbine efficiency and reliability. The SVMS methodology utilized advanced signal processing on vibration data captured in real time to perform condition-based monitoring and incipient fault detection. In [10], high frequency modal analysis, power trending and envelope analysis were used to identify

This work was supported in part by the US National Science Foundation (NSF) through grant no. ECCS-1809164 and an Early-Career Research Fellowship from the Gulf Research Program of the National Academies of Sciences, Engineering, and Medicine.

B. Freeman, Y. Tang and Y. Huang are with the Department of Computer & Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL 33431, USA (e-mail: {bfreem10, tangy, yhwang2018}@fau.edu).

J. VanZwieten is with the Department of Civil, Environmental, and Geomatics Engineering, Florida Atlantic University, Boca Raton, FL 33431, USA (e-mail: jvanzwi@fau.edu).

faulty bearings possessed by an emulated ocean current turbine (OCT).

The research performed in this work is focused on O&M cost reduction, as such costs were found to account for an estimated 26-32% of the LCOE [11]. Due to the lack of industry maturity, research relating to this field remains an under-represented area. Therefore, to address such headwinds, we have proposed an innovative framework that possess the following scientific merits:

- 1) The creation of a fault detection framework that does not require the use of a multi-sensor network. We propose a novel single sensor based methodology that only requires the use of a shaft key encoder.
- 2) To combat challenges associated with non-stationary signal analysis, our method incorporates the use of wavelet analysis to overcome issues related to spectral leakage.
- 3) Our framework incorporates the use of dimension reduction techniques for optimal fault feature representation learning. This step facilitates improved downstream machine learning-based fault detection and identification.

The remainder of this paper is structured as follows: Section II discusses the advancements made by related work in this field. Section III briefly references the MCT numerical simulation platform utilized in this work and touches upon the nature of the rotor blade pitch imbalance faults being analyzed. Section IV introduces our fault detection framework along with a background literature synopsis. Section V discusses our experimental design and quantifies its results. Lastly, Section VI presents the conclusion and future work.

II. RELATED WORK

Due to the intrinsic similarities between wind turbines and marine current turbines, many of the same fault detection methodologies developed for wind turbine generators (WTG) can easily be employed on MCT generators without major modification. Electrical stator current analysis is one of the most widely used fault detection techniques for such tasks [12]–[14], as it provides an optimal means of identifying abnormal frequency excitations within the characteristic frequency ranges of WTGs. Such methods traditionally incorporate the use of advanced signal processing and machine learning techniques due to their adeptness at facilitating varying degrees of non-intrusive condition based monitoring. However, since many of the signals utilized for non-intrusive fault detection are non-stationary in nature, the occurrence of spectral leakage adds an additional layer of complexity to their analysis. Spectral leakage occurs when the power of a signal is smeared across its frequency spectrum due to the signal not being periodic during some predefined sampling interval. This phenomenon adds additional frequency components to the spectrum of the signal, thereby masking important spectral details about the signal. To overcome such challenges, the following research has been carried out.

Synchronous Sampling-based Methods: Gong et-al developed a synchronous sampling algorithm for the detection of rotor eccentricities and bearing faults experienced by variable

speed direct drive wind turbines. The algorithm synchronously sampled the turbine's non-stationary generator stator current signal and applied an impulse detection module to identify fault signatures within the stator current's frequency spectrum. Unfortunately, this version of the algorithm is only applicable towards wind turbines possessing PMSGs, as frequency smearing is induced within the spectrum of the stator current signal when such a method is employed on a turbine possessing a DFIG [15]. In [15], Cheng et-al utilized synchronous sampling on the Hilbert Envelope obtained from the generator stator current signal of a DFIG wind turbine for gearbox fault detection. Since the stator current's envelope signal still retains fault frequencies proportional to the turbine's shaft rotating frequency, frequency excitations at the gearbox's characteristic frequencies were able to be easily identified when applying synchronous sampling on the amplitude demodulated envelope signal. However, such methods are limited to fault diagnosis, and provide no insight toward wind turbine fault prognosis.

Empirical Mode Decomposition-based Methods: In [16], Lu et-al leveraged the abilities of Empirical Mode Decomposition (EMD) and the Hilbert-Huang Transform (HHT) to extract wind turbine fault signatures related to rotor blade imbalances and inner-race way bearing faults from the frequency spectrum of the turbine's electrical stator current signal. Lu's proposed method accomplished this by accurately capturing the instantaneous amplitude and frequency of the generator current signals and then performing a comparative study between the signals in healthy and faulty conditions. In [13], Zhang utilized synchronous sampling and EMD in conjunction with a Generalized Likelihood Ratio (GLR) test to demodulated ocean current and turbulence fluctuation effects from the frequency spectrum of the generator's stator current signal. Doing so allowed for much easier fault signature identification within the frequency spectrum of the stator current signal. Challenges encountered when utilizing EMD-based methods for fault detection include the lack of a strong theoretical foundation, limitations associated with the creation of edge effects, strong reliances on the sifting of the stopping criteria, and extreme amounts of interpolation.

Wavelet-based Methods: In [17], Freeman employed the use of a Morlet Continuous Wavelet Transform (MCWT) on the generator power signal of a MCT. The wavelet coefficient energy contained within the frequency range of the turbine's rotor shaft (1P frequency) was quantified and utilized as a fault signature for rotor blade pitch imbalance faults. In [18], an adaptive Morlet wavelet was used to create a rolling bearing fault detection algorithm for wind turbine planetary gearboxes. Unfortunately, the performance of wavelet based fault detection lies within the user's ability to match the shape of the wavelet kernel to the shape of the fault signature within a signal of interest. Thus, it is imperative for the user to have an extensive domain knowledge in the wavelet field.

Artificial Intelligence-based Methods: In [19], a stacked autoencoder was used in conjunction with a support vector machine to extract fault features affecting gearboxes in DFIG wind turbines. In [20], an adaptive feature extraction algorithm that utilized signal resampling, frequency tracking, and a particle swarm optimized multi-class support vector machine

was developed to identify gearbox fault signatures masked within the stator current signal of a PMSG generator.

III. MCT MODELING AND IMBALANCE FAULTS ANALYSIS

A. MCT Modeling and Simulation

The MCT numerical simulation platform utilized for this work incorporates the blade element momentum modeling technique and dynamic wake inflow model presented in [21] to calculate the fluid structure interactions. Additionally, the 20-m diameter variable pitch rotor utilized by the platform is based upon the work done in [22]. The degrees of freedom of the rotor were limited to the rotation angle and velocity, and incorporates the “Tidal Turbine” version of the numerical simulation platform presented in [23]. The turbulence model accounts for spatial coherence over the swept area of the rotor bade as presented in [24], and a vertical shear profile that linearly decreases with depth. Including vertical shear was very important in this study as the coupling between the naturally occurring water shear and the rotor blade pitch imbalance is what is responsible for generating the primary fault signatures that appear in the MCT’s generated power signal. Thus, increasing the shear magnifies the fault signature while increasing the turbulence creates the fault signatures.

B. Analysis of Rotor Blade Pitch Imbalance Faults

Rotor blade pitch imbalance faults are induced by misaligned rotor blades. Such misalignments create differences in the vertical shear profiles experienced by the blades and have the potential to induce shaft torque variations [25].

The kinetic energy inherent within the varying dynamic loads and vibrations produced by shaft torque variations is transferred between the turbine’s rotor shaft and generator via the electromagnetic couplings shared between the two. It was shown in [26], that through such interconnections, amplitude and frequency modulations (AM and FM) are induced upon the generator’s electrical stator current signal. Thus, for a direct drive MCT under the influence of pitch imbalance faults, the shaft torque generated by the variable speed rotor can be described as:

$$T_t = T_{tw}(t) + A_{stv} \cdot \cos\left(\int 2\pi \cdot f_{st} dt\right) \quad (1)$$

where, T_t is the torque experienced by the rotor shaft, T_{tw} is the torque resulting from turbulence and wave activity, A_{stv} is the amplitude of the shaft torque variation stemming from the pitch imbalance fault, and f_{st} is the frequency of the shaft torque variation.

The resulting AM and FM affects both the generator stator current signal, $I_{gen}(t)$, and its fundamental frequency, $f_I(t)$, such that:

$$I_{gen}(t) = I_{vca}(t) + A_{pfc}(t) \cdot \sin\left(\int 2\pi \cdot f_{st} dt + \psi_{pf}\right) \quad (2)$$

$$f_{gen} = f_{vcf}(t) + A_{pff} \cdot \sin\left(\int 2\pi \cdot f_{st} dt + \varphi_{pf}\right) \quad (3)$$

where in (2), I_{gen} is the amplitude of the stator current signal, I_{vca} is the component of the stator current amplitude that is

generated from the variable ocean current power, A_{pfc} and ψ_{pf} are the respective amplitude and phase components of the signal that are created by the pitch imbalance fault. In (3), f_{gen} is the fundamental frequency of the stator current signal, f_{vcf} is the component of the fundamental frequency that is created by the variable ocean current power, and A_{pff} and φ_{pf} are the amplitude and phase components of the signal created by the pitch imbalance fault, respectively.

Combining equations (2) and (3) allows for the formulation of the modulated stator current signal:

$$C_{gen} = I_{gen}(t) \cdot \sin\left(\int 2\pi \cdot f_{gen} dt\right) \quad (4)$$

where C_{gen} is the AM and FM version of the generator stator current signal under pitch imbalance fault conditions.

Lastly, if it is assumed that the generator outputs an ideal three-phase supply voltage, then the instantaneous power can be described as:

$$P_{gen}(t) = V_{gen}(t) \cdot C_{gen}(t) \quad (5)$$

where $V_{gen}(t)$ is the single phase stator terminal voltage. Thus, the single-phase power of the generator can be described similarly to [27] as:

$$P_{gen}(t) = \frac{\sqrt{3}}{2} \cdot \left\{ \begin{aligned} &V_{Max} C_{Max} [\cos(2\omega_s t - \varphi_s) + \cos\varphi_s] \\ &+ \sum_{m=1}^{\infty} \left\{ \begin{aligned} &V_{Max} C_F \left[\cos((2\omega_s - m\omega_r)t - \varphi_F) \right. \\ &\quad \left. + \cos(m\omega_r t + \varphi_F) \right] \\ &+ V_{Max} C_{FN} \left[\cos((2\omega_s + m\omega_r)t - \varphi_{FN}) \right. \\ &\quad \left. + \cos(m\omega_r t - \varphi_{FN}) \right] \end{aligned} \right\} \end{aligned} \right\}$$

where V_{Max} is the maximum value of the supply line-to-line voltage, C_{Max} is the maximum value of the fundamental supply current, m is a positive constant integer, ω_s is the angular frequency of the stator’s supply current signal, and C_F is the maximum value of the characteristic fault component in the stator current signal at the frequency $f_{gen} + mf_{rf}$, where f_{rf} is the rotational frequency of the turbine’s rotor shaft. Additionally, φ_F is the initial value of the phase angle when the characteristic fault component has a frequency of $f_{gen} + mf_{rf}$, C_{FN} is the maximum value of the characteristic fault component in the stator current signal at the frequency $f_{gen} - mf_{rf}$, and φ_{FN} is the initial value of the phase angle when the characteristic fault component has a frequency of $f_{gen} - mf_{rf}$, and ω_r is the angular velocity of the turbine’s rotor.

This analysis yields the following conclusions: 1) Rotor blade pitch imbalance faults induce shaft torque variations; 2) The shaft torque variations create dynamic loads and vibrations that are transferred onto the rotor shaft [28]; 3) The kinetic energy housed within these dynamic loads and vibrations modulate the generator’s electrical power signal; and 4) The shaft torque variations create excitations within the frequency spectra of the electrical power signals. However, due to the induced AM, FM, and spectral leakage, these excitations become masked within the frequency spectrum of $P_{gen}(t)$. Thus, traditional Fourier based frequency analysis methods are unable to fully capture the entire range of such dynamics and highlights the need for more robust techniques.

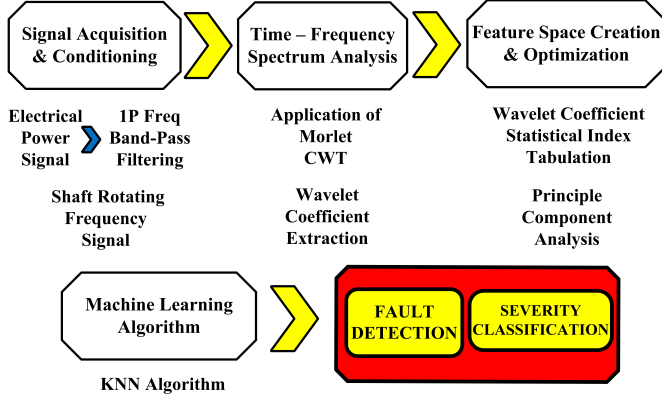


Fig. 1: Schematic diagram of the proposed fault detection and classification framework.

IV. PROPOSED FAULT DETECTION FRAMEWORK

Our novel fault detection and severity classification framework is proposed in Fig. 1, and begins with a signal acquisition phase, whereupon the instantaneous rotor frequency signal, f_{rot} , and $P_{gen}(t)$ are acquired from the turbine's generator. The signal f_{rot} is strictly used to determine the 1P frequency range that $P_{gen}(t)$ is band-passed filtered around and is not used afterwards.

Since the CWT is highly adept at analyzing the frequency spectrums of non-stationary signals, a time-frequency spectrum of $P_{gen}(t)$ is created utilizing a Morlet wavelet based CWT. The resulting wavelet coefficients are then tabulated into a six-degree feature space, where each dimension of the feature space corresponds to a different statistical index calculated from the wavelet coefficients. The chosen six statistical features are the mean, standard deviation, skewness, kurtosis, RMS, and peak to peak values.

Principle Component Analysis (PCA) is then employed to reduce the dimensionality of the feature space by only utilizing the components that correspond to the directions of maximum variation within the data set. Finally, a KNN machine learning algorithm is employed on the lower dimension feature space for the purpose of fault detection and severity classification. Additional details for each step of the framework are provided below.

A. Signal Acquisition and Conditioning

Band-pass filtering is performed on $P_{gen}(t)$ via the application of a Gaussian filter kernel that is mean centered at f_{rot} 's average frequency, such that the full width and half maximum of the kernel corresponds to the ± 3 standard deviation (3-STD) range of f_{rot} . This filtering removes the DC offset, sampling noise, and any other unrelated frequencies of interest from the the spectrum of $P_{gen}(t)$.

B. Time Frequency Spectrum Analysis

A Continuous Wavelet Transform (CWT), is used to generate the time-frequency spectrum of $P_{gen}(t)$. The CWT can be

interpreted as the convolution between $P_{gen}(t)$ and several localized wave-like functions of oscillatory nature. These wave-like functions possess finite energy, zero mean, and are all derived from a single mother wavelet basis function, $\Psi(t)$ [29], [30].

$\Psi(t)$ contains two hyper-parameters that can be tuned to increase its robustness for non-stationary signal analysis. The first parameter is the dilation parameter, $\mathbf{a}$, which controls the stretching and contraction of $\Psi(t)$. The second parameter is the translation parameter, $\mathbf{b}$, which shifts $\Psi(t)$ along the length of $P_{gen}(t)$. $\Psi(t)$ takes the general form: $\Psi(t) = \Psi((t - b)/a)$, utilizing a range of $\mathbf{a}$'s and $\mathbf{b}$'s. In the equation below:

$$\mathbf{T}(a, b) = w(a) \int_{-\infty}^{\infty} x(t) \Psi^* \left(\frac{t - b}{a} \right) dt \quad (6)$$

the quantities represented by $\mathbf{T}(a, b)$ are known as the wavelet coefficients, and are measures of cross-correlation between $P_{gen}(t)$ and $\Psi(t)$ [30]. Values of $\mathbf{T}(a, b)$ are usually visualized via a spectrogram image, whose axes represent the various translation and dilation parameters of $\Psi(t)$. Additionally, $w(a)$ is a weighting function that is customarily set equal to $1/\sqrt{a}$ to ensure that wavelets of the same scale all possess equal amounts of energy. Lastly, the $*$ symbol indicates that the complex conjugate of $\Psi(t)$ is used.

A wide variety of $\Psi(t)$ functions may be used when applying the CWT, with the optimal choice largely dependent upon the similarity of shape between $\Psi(t)$ and the fault signature being analyzed. Since this research is concerned with analyzing 1P band specific frequency activity, the increased control, precision, and shape of the Morlet wavelet's windowing kernel made it the ideal $\Psi(t)$ for this research [31].

A complex Morlet wavelet, w , can be constructed by multiplying a Gaussian window function with a sine wave: $w = e^{2i\pi f t} e^{(-t^2/2\sigma^2)}$, where i is the imaginary operator, f is the peak frequency in Hertz of the sine wave, and t is the time in seconds [32]. Additionally, $\sigma = n/2\pi f$, is the parameter that controls the width of the Gaussian window function, for which n dictates the trade off between frequency and time precision. Furthermore, n is an extremely non-trivial parameter that controls the Heisenberg uncertainty principle for time-frequency analysis and heavily influences the quality of results that are achievable from the data [31]. Through empirical analysis, n was selected to range from 5-15, with 100 increments between these values.

C. Feature Space Creation & Optimization

The wavelet coefficients generated by the CWT are statistically tabulated into a that is comprised of features corresponding to the mean, standard deviation, skewness, kurtosis, RMS, and peak to peak values. PCA is then utilized to reduce the dimensionality of this six-degree feature space by extracting out only the principle components associated with the directions of maximum variation within the data set.

PCA is a data reduction technique whose aim is to construct a set of principle components based upon the covariance amongst a set of correlated features in an N-dimensional data

set [31]. Each principle component represents a vector in N -dimensional space that characterizes the direction of a certain amount of variance contained within the data set.

Performing PCA begins with the construction of a covariance matrix, **CovMat** [31]:

$$\mathbf{CovMat} = (n - 1)^{-1} (\mathbf{X} - \bar{\mathbf{X}}) (\mathbf{X} - \bar{\mathbf{X}})^T \quad (7)$$

In the equation above, $\mathbf{X}$ is the given n by m data set or feature matrix being analyzed and $\bar{\mathbf{X}}$ is the mean value $\mathbf{X}$. After the covariance matrix has been calculated, an eigendecomposition of the covariance matrix is performed. Matlab's *eig* function makes this process convenient and efficient via the use of the functional equation: $[\mathbf{W}, \lambda] = \text{eig}(\mathbf{X})$, where $\mathbf{W}$ is an m by m matrix of the resulting principal components, and λ is a diagonal matrix of the associated eigenvalues. The eigenvectors characterize patterns within the covariance matrix, and can be viewed as a set of new coordinate axes to analyze the feature space with. The eigenvalues, λ , represent the magnitude of the eigenvectors.

Principle components are customarily ordered according to decreasing score value, T , such that the 1st principle component characterizes the "direction" of the largest amount of variance within the data set, the 2nd principle component characterizes the "direction" of the 2nd largest amount of variance within the data set, and so forth. Since $T = \mathbf{X} \cdot \mathbf{W}$, for which T is an n by m score matrix, is ordered, it can be truncated to contain only the " r " most relevant principle components, such that: $T_r = \mathbf{X} \cdot \mathbf{W}_r$. Here, $\mathbf{W}_r$ is an m by r matrix, and T_r is the n by r truncated score matrix obtained from T .

D. Application of Machine Learning Algorithm

A KNN machine learning algorithm is then employed on the truncated feature space for fault detection and severity classification.

The KNN machine learning algorithm is a lazy, non-parametric, supervised learning algorithm that is adept at solving both regression and classification problems. The decision making capability of the KNN algorithm is based on feature similarity, in that the algorithm assumes that similar instances exist in close proximity to each other. The prediction of a new instance is determined by a majority voting mechanism, in which the class of the new instance is chosen to be the same as that of the majority class of the K nearest instances (neighbors) around this new instance. The most common way to compute the distance between instances is via a Euclidean distance measure, $D = \sqrt{\sum_{i=1}^n (x_i - x_i')^2}$. When classifying a new instance, the selection of the most optimal amount of neighbors is usually done empirically.

V. EXPERIMENTS AND QUANTIFICATION OF RESULTS

A. Simulation Setup and Data Set Preparation

50 random seeds of data were simulated for the experiment, where for each seed a generator power and rotor frequency signal were simulated for each of the three studied fault cases (i.e., zero degrees, two degrees, and four degrees).

Each signal was sampled at a frequency of 10Hz and for a total duration of 300 seconds. The simulation platform utilized a turbulence intensity of 10% and a shear reduction rate of 0.0035 (m/s)/m with depth. The mean flow speed of the simulated ocean currents was 2 m/s at a max height of 20 m above the center of rotation of the rotor. To control the rotor speed, an industry standard fixed gain torque, $\tau = k \cdot \omega^2$, controller was used to maintain an operating speed near that of maximum power production. A representative look at the simulation output corresponding to seed one of the data set can be seen in Fig. 2.

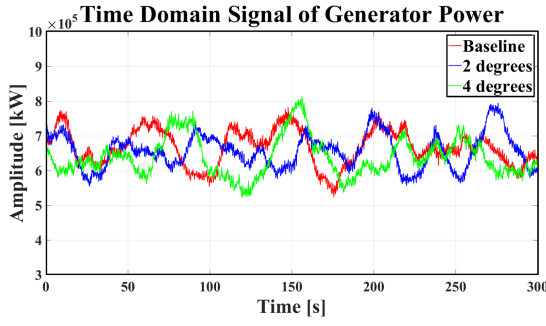
B. Experimental Results

Fig. 2a displays the $P_{gen}(t)$ signals corresponding to seed 1 of the data set. In Figs. 2b - d, the time-frequency spectrums of the $P_{gen}(t)$ signals are shown. The bottom portion of the spectrums portray the sum of the wavelet coefficient energy contained within the 1P frequency region, for which the center red line represent the mean of f_{rot} , and the two boundary red lines represents the ± 3 STD range limit of f_{rot} . While there does appear to be a clear distinction within the 1P band specific activity amongst these plots, establishing a consistent means of accurately quantifying the difference in this activity, throughout the entirety of the data set, is a fundamental goal of this work.

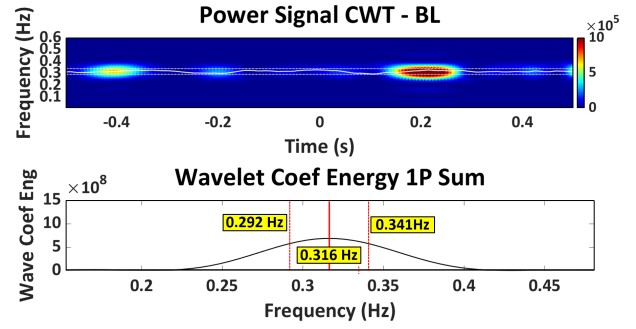
To aid in developing a means of accurately quantifying the 1P band specific frequency activity for all fault cases, a six dimensional statistical feature space was created to concisely describe the statistical characteristics of the wavelet coefficient energy contained within the ± 3 STD regions. Fig.3 is a plot of the Pearson's Correlation Matrix which is used to simultaneously investigation the linear dependencies between the different wavelet coefficient statistical features used to create the feature space.

Since the correlation coefficients possessed by many of the pairs of features have large magnitudes, PCA can be employed to reduce the dimensionality of the feature space. Figs.4a depicts the cumulative sum of variance possessed principle components, where it is shown that approximately 95% of the total variance contained within the feature space can be captured by these first two principle components. Figs.4b-c show the relative percentage wise contribution that each statistical features makes towards the creation of principle components one and two respectively. Furthermore, the shear and turbulence characteristics that are responsible for creating the analyzed fault signatures and their intensities in this research are captured very well via the use of PCA. As shown, principle component one has its largest contributions made by the RMS, Mean, and STD features, while principle component 2 is constructed primarily from the Skew and Kurt features.

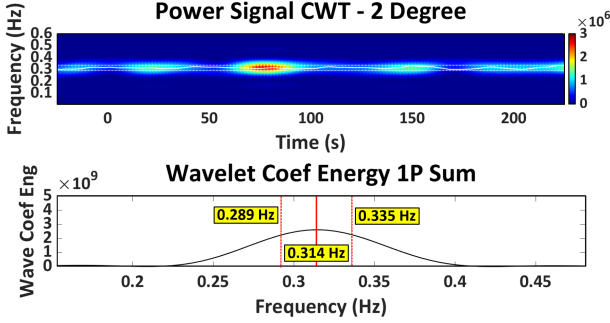
Fig.4d is the projection matrix obtained when the coordinate axes of the feature space are aligned to coincide with the directions of principle components one and two. To further increase the succinctness of the clustering within the figure, concentration ellipses are drawn around the barycenter of point belonging to each fault class. Discernible from the figure is the major benefit of using PCA, which lies within



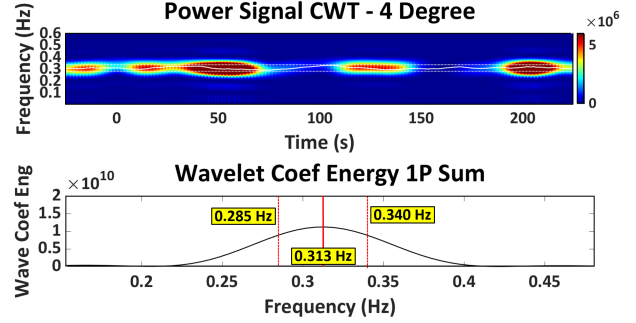
(a) Time series representation of the electrical power signal output by the turbine's generator



(b) CWT image of BL Power Signal



(c) CWT image of 2 Deg Power Signal



(d) CWT image of 4 Deg Power Signal

Fig. 2: (a) $P_{gen}(t)$ signals corresponding to seed 1 of the data set. (b) - (d) CWT Time-Frequency spectrums of the $P_{gen}(t)$ signals corresponding to (a). The bottom portion of the figures portray the sum of the wavelet coefficient energy contained within the 1P frequency region. The center red line represent the mean rotational frequency of the rotor, while the two boundary red lines represents the ± 3 STD range limits of the rotational frequency.

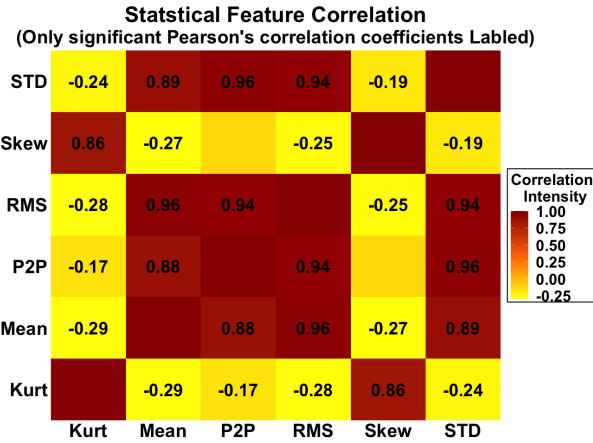


Fig. 3: Pearson's Correlation Matrix

the fact that instead of using a single or subset of features for classification, the principle components are in themselves linear combinations of all of the features used to create the feature space.

Before the KNN machine learning algorithm is employed the data set is partitioned into a 70% training set and 30% testing set. Fig.5a is a display of the projection matrix obtained after employing PCA on the testing partition. Repeated cross validation was utilized to train the KNN classifier, such that

10 fold cross validation was performed 10 distinct times on the training data set. The trained classifier was then employed on the testing data set displayed in Fig.5a. Correspondingly, Fig.5b displays the confusion matrix, for which the reported accuracy shows that the classifier is able to correctly predict the presence of a fault 100% of the time and correctly classify the severity of the fault 97.78% of the time. Through empirical analysis, the KNN algorithm was optimized to used five neighbors, for which a simple Euclidean distance measure was used to determine the distance between each neighbor.

Lastly, a summary table of the experimental results are presented in TABLE I. Of note in the table is the Kappa metric, which was found to be 96.67%. The Kappa metric is a measure of how closely the instances classified by the KNN classifier matched the ground truth while also taking into account the extent to which the data collected in the experiment is correctly represented by the measured features [33]. The Kappa metric is relevant because it removes the element of instances being correctly classified based upon random chance from its evaluation of accuracy. Additionally, the table also shows that the classifier proposed in our framework outperformed the No Information Rate, which is a metric produced when a naive classifier is employed with the purpose of proving that the model presented in this work possess a statistically significant P-value.

which compares the observed accuracy of the classifier with

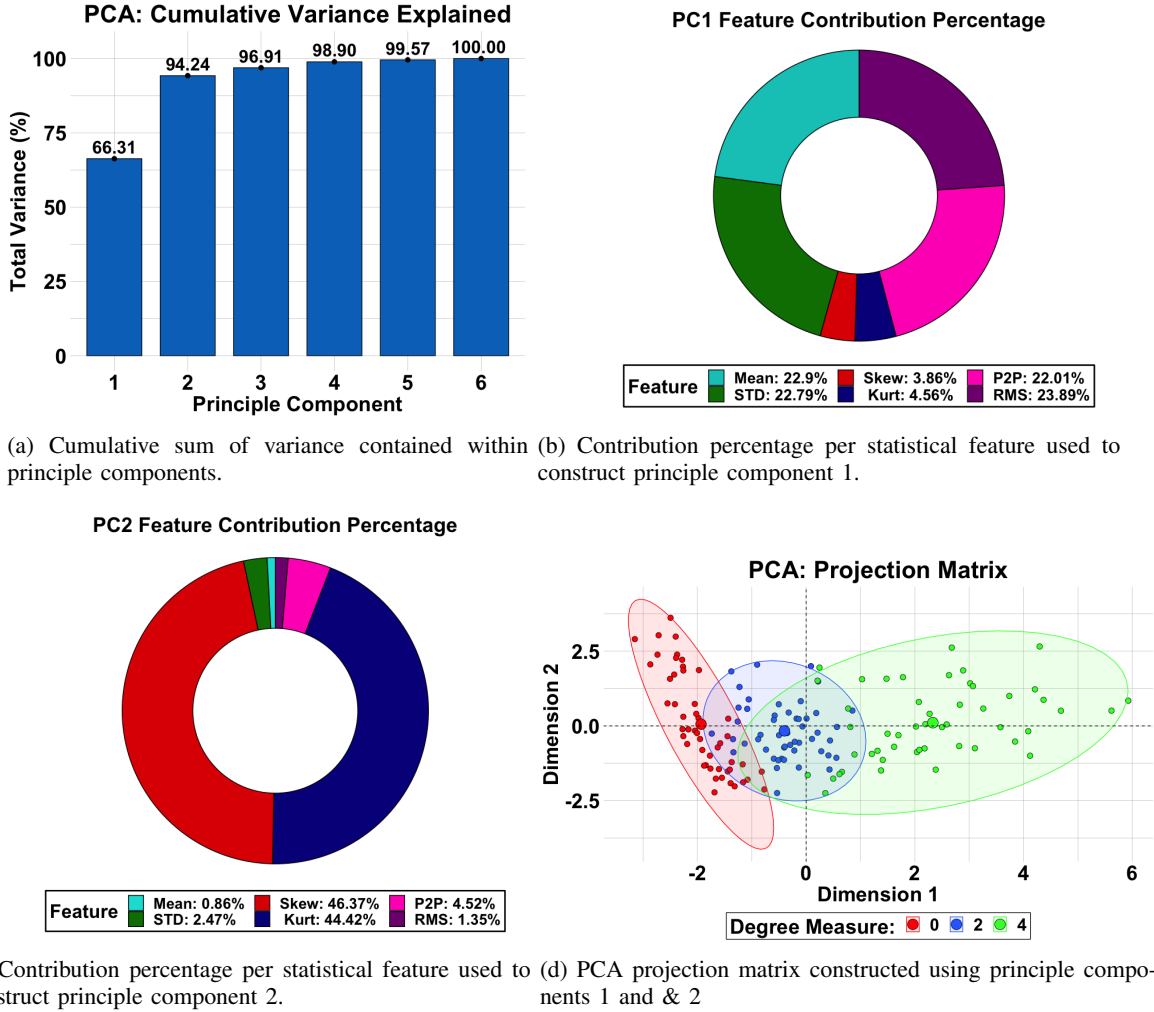


Fig. 4: (a) Cumulative Sum of Variance Explained per principle component plot. (b) Plot of the contribution of variance per statistical feature for principle component one. (c) Plot of the contribution of variance per statistical feature for principle component two. (d) Principle component projection matrix using only principle components one and two.

the expected accuracy of the classifier

Experimental Summary Statistics	
Kappa	0.9667
Accuracy	0.9778
95% CI Range	<0.8823 - 0.9994>
No Information Rate	0.3333
P - Value (Accuracy >NIR)	2.2e-16

TABLE I: Summary of Experimental Results

VI. CONCLUSIONS

The objective of this research was to develop a fault detection and severity classification framework for MCT rotor blade pitch imbalance faults. In the proposed framework, a Morlet CWT was first utilized to view the time-frequency spectra of the turbine's electrical power signals. Next, the wavelet coefficient energy encompassed within the 1P rotational frequency range of the turbine's rotor shaft were extracted and

statistically tabulated into a six-degree feature space. PCA was then used to reduced the dimensionality of the feature space so that the new coordinate axes of the space are aligned with the directions of maximum variation. Lastly, a KNN machine learning algorithm was employed on the resulting feature space for fault detection and severity classification. This research found that the KNN machine learning algorithm correctly predicted the presence of pitch imbalance faults 100% of the time and correctly classify their severity 97.78% of the time.

Our proposed method has very low hardware cost, and does not require the use of complex sensor networks (containing vibration, strain, torque, and/or acoustic emission sensors) that are used in contemporary fault detection and condition monitoring systems. Additionally, the proposed framework can also be effortlessly integrated into existing MCT control systems. In the future, it is hoped that the robustness of this framework can be expanded to allow for the prediction and classification of an extended range of faults that affect MCT systems, such as gear faults, drivetrain and bearing faults, and

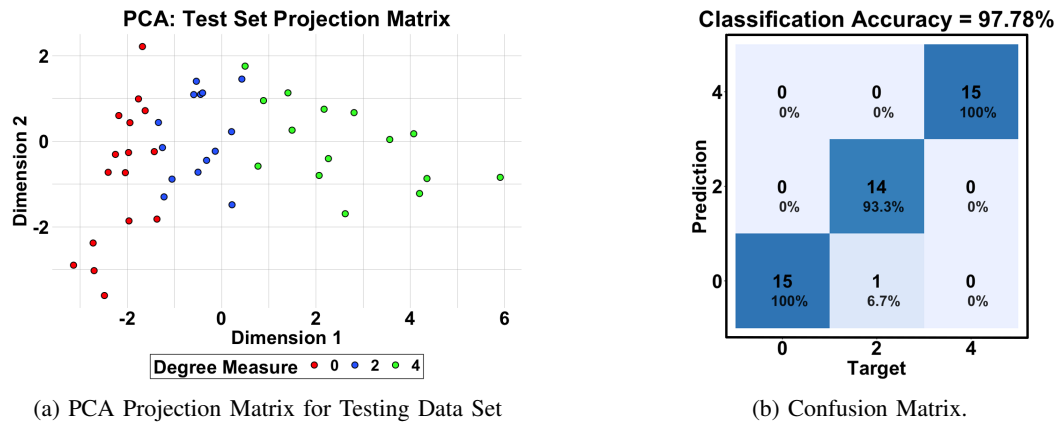


Fig. 5: (a) Projection Matrix for testing partition of the data set. (b) Confusion matrix results after the KNN machine learning algorithm was applied on the feature space created by the PCA projection Matrix in (a).

even sensor faults.

REFERENCES

- [1] K. Haas, "Assessment of energy production potential from ocean currents along the united states coastline," Georgia Tech Research Corporation, Tech. Rep., 2013.
- [2] M. C. P. Machado, J. H. VanZwieten, A. P. Callou, I. Pinos *et al.*, "A measurement based analyses of the hydrokinetic energy in the gulf stream," vol. 3, no. 1. IEEE, 2016, pp. 25–30.
- [3] P. Jacobson, "Assessment and mapping of the riverine hydrokinetic energy resource in the continental united states," *Electric Power Research Institute (EPRI)*, 2012.
- [4] G. T. Research, "Assessment of energy production potential from ocean currents along the united states coastline," 2013.
- [5] A. Felix, J. V. Hernández-Fontes, D. Lithgow, E. Mendoza, G. Posada, M. Ring, and R. Silva, "Wave energy in tropical regions: Deployment challenges, environmental and social perspectives," *Journal of Marine Science and Engineering*, vol. 7, no. 7, p. 219, 2019.
- [6] A. LiVecchi, A. Copping, D. Jenne, A. Gorton, R. Preus, G. Gill, R. Robichaud, R. Green, S. Geerlofs, S. Gore *et al.*, "Powering the blue economy; exploring opportunities for marine renewable energy in maritime markets," *US Department of Energy, Office of Energy Efficiency and Renewable Energy*. Washington, DC, p. 207, 2019.
- [7] Z. Zhou, F. Scullier, J. F. Charpentier, M. E. H. Benbouzid, and T. Tang, "Power control of a nonpitchable pmsg-based marine current turbine at overrated current speed with flux-weakening strategy," *IEEE Journal of Oceanic Engineering*, vol. 40, no. 3, pp. 536–545, 2015.
- [8] Y. Tang, Y. Zhang, A. Hasankhani, and J. VanZwieten, "Adaptive super-twisting sliding mode control for ocean current turbine-driven permanent magnet synchronous generator," in *2020 American Control Conference (ACC)*, 2020 (Accepted).
- [9] M. Mjit, P.-P. J. Beaujean, and D. J. Vendittis, "Smart vibration monitoring system for an ocean turbine," in *2011 IEEE 13th International Symposium on High-Assurance Systems Engineering*. IEEE, 2011, pp. 252–260.
- [10] N. Waters, P.-P. Beaujean, and D. Vendittis, "Detection, localization, and identification of bearing faults on an ocean turbine," in *2012 Oceans*. IEEE, 2012, pp. 1–8.
- [11] V. S. Neary, M. Lawson, M. Previsic, A. Copping, K. C. Hallett, A. LaBonte, J. Rieks, and D. Murray, "Methodology for design and economic analysis of marine energy conversion (mec) technologies." Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), Tech. Rep., 2014.
- [12] M. Zhang, T. Wang, T. Tang, M. Benbouzid, and D. Diallo, "Imbalance fault detection of marine current turbine under condition of wave and turbulence," in *IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society*. IEEE, 2016, pp. 6353–6358.
- [13] —, "An imbalance fault detection method based on data normalization and emd for marine current turbines," *ISA transactions*, vol. 68, pp. 302–312, 2017.
- [14] F. Mekri, S. Ben Elghali, and M. E. H. Benbouzid, "Fault-tolerant control performance comparison of three- and five-phase pmsg for marine current turbine applications," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 2, pp. 425–433, April 2013.
- [15] F. Cheng, L. Qu, W. Qiao, C. Wei, and L. Hao, "Fault diagnosis of wind turbine gearboxes based on dfig stator current envelope analysis," *IEEE Transactions on Sustainable Energy*, 2018.
- [16] D. Lu, W. Qiao, X. Gong, and L. Qu, "Current-based fault detection for wind turbine systems via hilbert-huang transform," in *2013 IEEE Power & Energy Society General Meeting*. IEEE, 2013, pp. 1–5.
- [17] B. Freeman, Y. Tang, and J. VanZwieten, "Marine hydrokinetic turbine blade fault signature analysis using continuous wavelet transform," in *2019 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2019, pp. 1–5.
- [18] Y. Kong, T. Wang, Z. Li, and F. Chu, "Fault feature extraction of planet gear in wind turbine gearbox based on spectral kurtosis and time wavelet energy spectrum," *Frontiers of Mechanical Engineering*, vol. 12, no. 3, pp. 406–419, 2017.
- [19] F. Cheng, J. Wang, L. Qu, and W. Qiao, "Rotor-current-based fault diagnosis for dfig wind turbine drivetrain gearboxes using frequency analysis and a deep classifier," *IEEE Transactions on Industry Applications*, vol. 54, no. 2, pp. 1062–1071, March 2018.
- [20] D. Lu and W. Qiao, "Fault diagnosis for drivetrain gearboxes using pso-optimized multiclass svm classifier," in *2014 IEEE PES General Meeting | Conference Exposition*, July 2014, pp. 1–5.
- [21] J. H. VanZwieten, N. Vanrietvelde, and B. L. Hacker, "Numerical simulation of an experimental ocean current turbine," *IEEE Journal of Oceanic Engineering*, vol. 38, no. 1, pp. 131–143, Jan 2013.
- [22] J. H. VanZwieten, P. Pyakurel, T. Ngo, C. Sultan, and N. I. Xiros, "An assessment of using variable blade pitch for moored ocean current turbine flight control," *International journal of marine energy*, vol. 13, pp. 16–26, 2016.
- [23] V. Tzelepis, J. H. VanZwieten, N. I. Xiros, and C. Sultan, "System modeling and simulation of in-stream hydrokinetic turbines for power management and control," *Journal of Dynamic Systems, Measurement, and Control*, vol. 139, no. 5, 2017.
- [24] P. Pyakurel, J. H. VanZwieten, M. Dhanak, and N. I. Xiros, "Numerical modeling of turbulence and its effect on ocean current turbines," *International journal of marine energy*, vol. 17, pp. 84–97, 2017.
- [25] X. Gong and W. Qiao, "Imbalance fault detection of direct-drive wind turbines using generator current signals," *IEEE Transactions on energy conversion*, vol. 27, no. 2, pp. 468–476, 2012.
- [26] —, "Bearing fault diagnosis for direct-drive wind turbines via current-demodulated signals," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 8, pp. 3419–3428, 2013.
- [27] S. J. Watson, B. J. Xiang, W. Yang, P. J. Tavner, and C. J. Crabtree, "Condition monitoring of the power output of wind turbine generators using wavelets," *IEEE Transactions on Energy Conversion*, vol. 25, no. 3, pp. 715–721, 2010.
- [28] D. Jiang, Q. Huang, and L. Hong, "Theoretical and experimental study on wind wheel unbalance for a wind turbine," in *2009 World Non-Grid-Connected Wind Power and Energy Conference*, Sept 2009.
- [29] R. Yan, R. X. Gao, and X. Chen, "Wavelets for fault diagnosis of rotary

- machines: A review with applications,” *Signal Processing*, vol. 96, pp. 1–15, 2014.
- [30] P. S. Addison, *The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance*. CRC press, 2017.
- [31] M. X. Cohen, *Analyzing neural time series data: theory and practice*. MIT press, 2014.
- [32] —, “A better way to define and describe morlet wavelets for time-frequency analysis,” *bioRxiv*, 2018. [Online]. Available: <https://www.biorxiv.org/content/early/2018/08/21/397182>
- [33] M. L. McHugh, “Interrater reliability: the kappa statistic,” *Biochemia medica: Biochemia medica*, vol. 22, no. 3, pp. 276–282, 2012.